SAND2013-2542J

A Bayesian method for using simulator data to enhance human error probabilities assigned by existing HRA methods

Katrina M. Groth[a], Curtis L. Smith[b], Susan M. Stevens-Adams[a], Laura P. Swiler[a]

[a] *Sandia National Laboratories, Albuquerque, NM 87185-0748, USA*
[b] *Idaho National Laboratory, Idaho Falls, ID 83415-3850, USA*

Abstract

In the past several years, several international agencies have begun to collect data on human performance in nuclear power plant simulators [1]. This data provides a valuable opportunity to improve human reliability analysis (HRA), but there improvements will not be realized without implementation of Bayesian methods. Bayesian methods are widely used in to incorporate sparse data into models in many parts of probabilistic risk assessment (PRA), but Bayesian methods have not been adopted by the HRA community. In this paper, we provide a Bayesian methodology to formally use simulator data to refine the human error probabilities (HEPs) assigned by existing HRA methods. We demonstrate the methodology with a case study, wherein we use simulator data from the Halden Reactor Project to update the probability assignments from the SPAR-H method. The case study demonstrates the ability to use performance data, even sparse data, to improve existing HRA methods. Furthermore, this paper also serves as a demonstration of the value of Bayesian methods to improve the technical basis of HRA.

Keywords: Human Reliability Analysis (HRA), Bayesian inference, simulator data, nuclear power plant, operator performance data

1. Introduction

Probabilistic risk assessment (PRA) is a conceptual and computational framework for making robust, defensible decisions in many industries, including nuclear power production. The goal of PRA is to estimate the both the probability and the consequences of accidents by systematically incorporating information from many sources. In making decisions about potentially hazardous systems, it is important to use all available information, including data and subjective information. As Aven states, “the purpose of risk analysis is to support decision-making, not to produce numbers” [2]. With this goal in mind, all information (objective and subjective) becomes relevant to making good decisions.

Bayesian methods are the sole framework that systematically incorporates subjective information into logical inference. Because of PRA’s focus on low-frequency scenarios in highly reliability systems, empirical data are often lacking. Consequently, Bayesian methods are widely used in PRA for hardware failure quantification [3, 4].

The hardware PRA community has built a framework for formally assigning probabilities using both subjective information and objective data [3, 4, 5, 6]. Within this framework, operational experience and data play an important role in the assignment of probabilities. This use of Bayesian methods can be complementary to the use of Bayesian Networks or can apply outside of the BN framework, as is done for hardware PRA.

There have been repeated calls to expanding the technical basis of Human Reliability Analysis (HRA) by systematically integrating information from different domains [7, 8, 9, 10, 11, 12]. Paradoxically, the HRA community has not embraced Bayesian methods, despite the fact that HRA has a greater amount of subjective information and informal evidence than the hardware community. The one exception has been in the area of causal models (called Bayesian Networks, BNs, or Bayesian Belief Networks, BBNs), which are becoming more prevalent among HRA researchers [13, 14, 15, 16, 17, 18, 19, 20, 21, 22] This BN research provides a powerful opportunity to enhance HRA, but BNs are only one implementation of the larger world of Bayesian methods that can benefit HRA. There remains a full range of Bayesian methods and associated computational techniques that can be applied within HRA, but which the HRA community has not adopted.

The use of Bayesian methods for refining various parts of HRA has been explored as an academic activity in NUREG/CR-6949 [23], but has not been adopted in the larger HRA community. In the time since NUREG/CR-6949, discussions about the use of data in HRA have continued to focused on using only "objective" frequencies (e.g., [24]), despite the fact that no single source of data can provide such frequencies. This narrow focus on frequencies is inconsistent with the rest of PRA (which encompasses subjective information), not to mention an unrealistic expectation for data collection. Furthermore, this objective-frequency approach discards substantial amounts of subjective data that is relevant to the decision making process. To enable the use of data and to ensure compatibility with the remainder of PRA, the HRA community must embrace an expanded range of Bayesian methods.

Using Bayesian methods entails three things: 1) using subjective probabilities[1] to express degrees of belief; 2) using all available information (subjective and objective) to assign these probabilities; and 3) applying the Bayesian inference process to update beliefs as additional information becomes available. The Bayesian inference process entails generating a hypothesis, describing knowledge about the hypotheses by use of a probability distribution (called the prior), collecting and filtering data, and then obtaining the overall probability distribution (called the posterior) which represents degree of belief in the hypothesis, conditional on all available evidence and information. The Bayesian inference process can be performed implicitly by experts (e.g., as a mental calculation) or with formal, explicit mathematical techniques (as is often done in PRA).

Conceptually, Bayesian methods can be applied to a wide range of HRA problems. Applying these methods computationally requires a more detailed look at specific problem spaces. NUREG/CR-6949 has been criticized as too general, and reviewers requested more targeted applications of Bayesian methods to specific problems [23, p. 59-60]. One such application is the use of Bayesian methods to formally incorporate data into the models that assign probabilities.

In this paper, we provide a methodology for using Bayesian methods and computational techniques traditionally used in hardware PRA to formally use simulator data to refine the Human Error Probabilities (HEPs) from an existing HRA method. We demonstrate the methodology by using simulator data from the Halden Reactor Project to update the SPAR-H HRA method (the prior). The results (posterior) discussed in this paper are specific to this combination of information, but the methodology provided is applicable to any combination of HRA methods and data. The intention is twofold: to demonstrate the expanded use of Bayesian methods within HRA, and to demonstrate that it is possible to use data, even sparse data from different sources, to update existing HRA methods.

In Section 2, we argue that parameters of existing HRA models can be interpreted as prior parameters, and that they should be subject to Bayesian updating in the same way that the parameters of an aleatory model are updated with hardware data. In Section 3 we provide an introduction to the Bayesian inference process. In Section 4, we present the methodology for using simulator data to update beliefs about human error probability. In Section 5, we present a case study where we used data extracted from Halden simulator experiments to update HEPs from the SPAR-H method. In Section 6, we provide a discussion of the work, and we present final conclusions in Section 7.

2. The meaning of probability in PRA and HRA

PRA uses a combination of models to predict system behavior and possible consequences and to assign probabilities to different contributors and outcomes. Probability is used as a tool to distinguish more likely scenarios from less likely scenarios. The scenarios are decomposed into a set of basic events (e.g., component failure on demand, human errors), which are encoded in deterministic logic-based models (e.g., Fault Trees and Event Trees). A probability is associated with each basic event, and the basic event probabilities are com-

[1]Objective or empirical data can be used to develop subjective probabilities, but the use of the subjective interpretation or probability also allows additional types of information to be used.

bined according to the logic model to provide scenario probabilities for use in decision making.

An example basic event in PRA is "failure of valve in a closed position." The probability of each basic event is expressed by a probability model and an associated set of parameters. Each probability model could be a common probability distribution, such as the Bernoulli distribution with parameter p or the binomial distribution with parameters p and n. Each could also be a more complicated probability model such as an influence diagram, which incorporates causal factors, each of which is associated with a set of parameters. For either modeling approach the analyst assigns model parameters, which represent his/her current state of knowledge based on all of the available information. These models are derived from either (a) subjective information (e.g., about the chance of not performing an intended function), (b) actual failure data, or a combination of (a) and (b).

The goal of quantitative HRA is to determine the human error probability (HEP), which is the probability of the basic event associated with a human failure event (HFE). Example events in PRA are "failure to initiate Feed and Bleed" and "failure to align electrical bus to alternative feed." In probability language, the HEP is an assignment of the belief of the truth of the proposition "the human response to event X will not satisfy system requirement Y." By replacing the word "human" in the previous sentence with the word "hardware," it becomes apparent that assigning the HEP with an HRA model is not conceptually different than assigning a hardware failure probability with a common probability distribution such as the binomial distribution.

The difference between HRA and hardware analysis is that HRA uses HRA models/methods to assign probabilities, whereas for hardware, statistical data is used in combination with subjective information to assign probabilities. There are numerous HRA methods available that provide models for assigning the HEP. Many HRA methods provide a function that assigns the HEP, based on the context of the performance. That is, these provide the conditional probability of a human failure event (HFE), given the context of performance $P(HFE|context)$. In many HRA methods, the context is represented by a set of Performance Shaping Factors (PSFs) or Performance Influencing Factors (PIFs), which are discretized into levels or states.

The use of detailed HRA models to predict outcomes for human events is no different than the use of a binomial distribution to predict outcomes for hardware events. The only difference between the two cases is the number of parameters included in the model used to make the prediction.

3. Overview of Bayesian inference

The goal of applying Bayesian inference is to reason about a set of hypotheses. In other words, we will determine which hypotheses are logically more plausible, given the initial knowledge and the evidence/data available. The hypothesis, H, expresses a belief, e.g., about an event or a set of outcomes or about the parameters of a model. The data, D, represents observations relevant to the hypothesis.

The heart of the Bayesian inference process is Bayes' Theorem (Equation 1), which provides the conceptual and mathematical means for combining different information in the context of a probabilistic model.

$$P(H|D) = P(H)\frac{P(D|H)}{P(D)} \tag{1}$$

We apply Bayes' Theorem to obtain the posterior distribution, $P(H|D)$, which expresses the plausibility of the hypothesis, given the data. We follow this process to determine the probability that a hypothesis is true, conditional on all available evidence. Bayes' Theorem can be applied iteratively to incorporate a variety of new evidence to make inferences about the same hypothesis. As such, the posterior probability from one analysis becomes the prior probability for the next. This process allows us to coherently incorporate a variety of information sources into a single structure in order to make inferences.

To apply Bayes' Theorem, information known about a hypothesis (independent of the data) is described with a prior; this prior is encoded via a probability model, $P(H)$. The prior is a probability distribution which expresses the plausibility of the hypothesis; this prior represents the analyst's initial knowledge about the hypothesis[2].

The next step is to collect or obtain data, D, and express it in a format that can be used to update the

[2] To reiterate, the Bayes' prior is intended to capture a state of knowledge independent of data collection. This does not imply that the event we are modeling has this probability distribution as an inherent property; rather, the probability distribution expresses beliefs about the event.

prior distribution. To do this, we use a likelihood model. The likelihood model allows us to translate the observable data into probabilistic information. The likelihood model represents the process providing the data; in other words, it is a model of performance. This likelihood expresses the probability of the data, given the truth of the hypothesis, $P(D|H)$. In Equation 1 this likelihood is normalized by the probability of the data over all specified hypotheses, $P(D)$. For k hypotheses, $P(D)$ can be calculated as $P(D) = \sum_{i=1}^{k} P(H_i)P(D|H_i)$. Different likelihood models will be appropriate for different types of data and hypotheses.

Bayes' Theorem provides the mathematical framework for assessing the posterior distribution, but executing the calculations can be an involved process. For discrete probability distributions, the posterior distribution can be easily calculated using Equation 1. However, for continuous probability distributions, the calculation of $P(D)$ involves integration, as shown in the continuous form of Bayes' Theorem (Equation 2).

$$\pi_1(\theta|x) = \frac{f(x|\theta)\pi_0(\theta)}{f(x)} = \frac{f(x|\theta)\pi_0(\theta)}{\int f(x|\theta)\pi_0(\theta)d\theta} \quad (2)$$

This integration cannot be obtained in a closed form for many forms of the likelihood function. In these cases, sampling methods such as Markov Chain Monte Carlo Methods (MCMC) can be used to solve for posterior distributions. MCMC sampling can be implemented in programs such as OpenBUGS (formerly WinBUGS) or Matlab. For some combinations of likelihood functions and prior distributions that are *conjugates*, the integration has a closed form. For these conjugate distributions, the prior and the posterior distribution are of the same functional type, and the posterior distribution can be calculated directly. Many of likelihood functions commonly used in PRA have conjugate prior distributions [6].

3.1. Example application in PRA

In PRA, analysts are tasked with expressing a probabilistic model for each basic event. One of the most common types of basic event in PRA is "component fails on demand." For these events and other Bernoulli processes, the most commonly used probability model is the binomial distribution, (Equation 3).

$$Pr(X = x) = f(x|p) = \binom{n}{x} p^x (1-p)^{n-x} \quad (3)$$

The binomial distribution is used to describe uncertainty about the number of failures, x, that will occur in a given number of demands, n, given a parameter, p, which is interpreted as the probability of failure-on-demand for the component. In PRA, this type of uncertainty is called aleatory uncertainty. The term aleatory refers to the stochastic or random nature of the outcome of processes such as coin flipping or valve opening.

The parameter p is not directly observable, but it can be inferred from data or assigned by experts. In most PRA applications the value of the parameter is uncertain, due to sparse data, partially relevant data, imprecision in the data or the methods, and so on. This type of uncertainty is denoted epistemic uncertainty. In PRA, the term epistemic refers to lack of knowledge about models and parameters. In PRA, the Bayesian inference process is used to express beliefs about the plausibility of different possible values for the parameter (that is, to provide the plausibility of hypotheses of the form "the value of parameter p is z").

Prior beliefs about the parameter can be specified using a variety of distributions, depending on the information available. The beta distribution is commonly used as prior distribution for the parameter, p, of the binomial distribution: $\pi(p|\alpha,\beta) \sim \mathrm{Beta}(\alpha,\beta)$. $\mathrm{Beta}(\alpha,\beta)$ indicates a beta random variable, which has probability density function (pdf) given by Equation 4.

$$f(p;\alpha,\beta) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)} \quad (4)$$

Where the beta function, $B(\alpha,\beta)$, is used as a normalization constant.

The beta distribution has the advantage of being conjugate to the binomial distribution. This means that the posterior distribution for this combination will also be a beta distribution, with parameters α_{post} and β_{post}. The values of these parameters, which represent the combination of the prior and the data, are given by Equation 5.

$$\begin{aligned} \alpha_{post} &= \alpha_{prior} + x \\ \beta_{post} &= \beta_{prior} + n - x \end{aligned} \quad (5)$$

Since the beta and binomial distributions are conjugate, the Bayesian updating process is simple: the analyst sets the prior belief about p by assigning values to α and β. The data records x failures in n demands. The posterior distribution for p is a beta distribution with parameters obtained from Equation 5. This posterior model expresses epistemic uncertainty about the value of p, which is a parameter of the aleatory model (Equation 3) to be used in the PRA. To implement the posterior distribution within the aleatory model, several approaches could be used. Commonly approaches include sampling from the distribution for p and using either the moments (e.g., the mean) or the percentiles of the distribution for p.

4. Methodology to Bayesian update HEPs with simulator data

Once we understand that assigning HEPs is not conceptually different than assigning hardware event probabilities, the Bayesian inference process can be applied to update HEPs. Executing the Bayesian inference process outlined in Section 3 requires several steps:

1. Define the hypothesis
2. Identify suitable sources of information
3. Specify the prior distribution
4. Specify the likelihood function
5. Conduct Bayesian updating to obtain a posterior

In this section, we demonstrate how to conduct the Bayesian updating process on HEPs generated by existing HRA methods using simulator data. The prior distribution is based on a current HRA method, and the likelihood function will be specified to match simulator data. It should also be noted that we could combine two or more priors to have a mixture prior representing multiple methods, however this discussion is beyond the scope of this paper.

4.1. Step 1: Define the hypothesis

The goal of HRA is to assign a human error probability (HEP) for the human failure event (HFE) X. In PRA language, we want to assign a probability the basic event "failure of human response to event X," which is assumed to be Bernoulli process. As in Section 3.1, we use the binomial distribution (with unknown parameter p and $n = 1$) as the aleatory model. Since we are uncertain about the value of p, we wish to use Bayesian inference to obtain a posterior probability distribution for p. Therefore, the goal of the Bayesian inference is to express a belief about the hypothesis "the HEP for event X is p."

4.2. Step 2: Identify suitable sources of information

The HRA method and the simulator data selected to be used in the Bayesian updating process must each be compatible with the hypothesis described in Step 1. Any HRA method that directly assigns human error probabilities is a candidate for use in Bayesian inference on the hypothesis. This includes methods such as ASEP [25], CBDT [26], CESA-Q [27], CREAM [28], HEART [29], KHRA [30], SLIM-MAUD [31], SPAR-H [32], THERP [33], and others. While the focus of the current paper is on HRA methods that directly assign HEPs, the same Bayesian updating process can be used to update expert judgments (such as those used to assign the probabilities as part of methods like ATHEANA [34]).

There are several international projects collecting data on human performance data in nuclear power plant simulators, including projects at the OECD Halden Reactor Project [35, 36], KAERI [37, 38], the U.S. NRC [39, 24] and others [1]. Each of these simulator data sources should be compatible with the hypothesis, although many of the data sources are not yet available to the HRA community. Data availability remains a key issue in HRA [10], so the choice of simulator data depends largely on availability.

In addition to being compatible with the hypothesis, the HRA method and the simulator data must be compatible with each other. Event decomposition must be consistent between the HRA method and the simulator data (i.e., they must both define the human failure event in a similar way.). The variables (e.g., PSFs or PIFs) in the data source must be defined in such a way that enables them to be mapped onto the variables in the HRA method. Alternatively, both the data and the HRA method can be mapped to the generic, comprehensive set of PIFs developed to enable data analysis [18].

4.3. Step 3: Specify the prior distribution

The prior distribution encodes information related to our knowledge about the hypothesis identified in Step 1. The form of the prior distribution depends on the quality of the prior information. When prior information is absent, it is desirable to use a non-informative prior to represent ignorance. When more information is available, informative prior distributions can be specified.

Existing HRA methods provide information, which can be used to express the prior distribution on p, which is denoted p_0. Most HRA methods provide a point estimate, or a formula to obtain a point estimate, of the value of p for each context or combination of PSFs. This point estimate can be interpreted as the mean value of p_0, $E(p_0)$ for a given context. To implement this in a probability distribution we use a limited information prior, based on the constrained non-informative (CNI) distribution, where the constraint is on the mean value of the distribution [40]. The CNI distribution can be approximated by a Beta(α, β) distribution, with $\alpha = 0.5$, and β derived from the constraint on the mean[3] given by Equation 6.

$$E(Beta(\alpha, \beta)) = \frac{\alpha}{\alpha + \beta} = E(p_0) \tag{6}$$

In PRA, α can be thought of as the number of failures contained in the prior distribution, and the sum $(\alpha + \beta)$ can be thought of as the number of demands over which these failures occurred. Therefore the CNI prior represents half of a failure occurring in $(\alpha + \beta)$ demands, where β is determined from the predictions of the HRA model. We use this procedure to develop a CNI prior for each context represented in the HRA method.

4.4. Step 4: Specify the likelihood function

The likelihood is a mathematical function representing the process providing the data. In a Bayesian analysis, the likelihood is conditioned on the data. Thus, the form of the likelihood must coincide with the type of data to be collected. For the hypothesis indicated in Step 1, we are reasoning about the likelihood of observing a human failure upon demand. For this problem, the relevant data is the number of human failures, x, in a given the number of demands or opportunities for human failure, n. The binomial likelihood function is appropriate for this type of data.

This likelihood model does not explicitly factor in information about the context of the performance (i.e., the PSFs). Rather, a different likelihood model is developed for each combination of PSFs in the HRA method.

4.5. Step 5: Conduct Bayesian updating to obtain a posterior

In the previous four steps, we established a hypothesis, gathered data relevant to the hypothesis, expressed prior beliefs about the hypothesis, and identified the appropriate likelihood model. In this step, we implement Bayes' Theorem to quantify the posterior belief about the hypothesis. In other words, we will determine which the plausibility of the hypothesis given the prior and the data.

Since we used a conjugate prior distribution and likelihood function, the Bayesian updating process is straightforward. The posterior distribution for p, denoted p_1 is Beta(α_{post}, β_{post}), where the values of α_{post} and β_{post} are established using Equation 5.

5. Case study: SPAR-H with Halden data

In this section, we implement the methodology detailed in Section 4 to perform updating on the SPAR-H method using simulator data from the Halden Reactor Project.

5.1. SPAR-H

The SPAR-H [32] human reliability analysis method was developed to estimate HEPs for use in the SPAR nuclear power plant PRA models. SPAR-H is used as part of PRA in over 70 U.S. nuclear power plants and by the event assessment programs at the NRC.

The SPAR-H method considers two plant states: at-power and low power/shutdown, and two types of human activities: diagnosis and action. The two types of activities use the same equations and PSFs, but use different PSF multipliers and different values for the nominal HEP (NHEP). In this paper, we present the model for action tasks during at-power operations.

The SPAR-H method uses eight PSFs to represent the context. Each PSF level is associated with an HEP multiplier value. Table 1 contains the SPAR-H PSFs and the PSF multiplier values for action tasks[4]. The first step in applying SPAR-H is to evaluate the level for each PSF to determine the multipliers.

The second step is to calculate HEP using equations provided in the worksheets. Two equations are provided; which equation is used depends on the

[3] This relationship is used when $E(p_0) \leq 0.5$. For $E(p_0) \geq 0.5$, the values are switched: $\beta = 0.5$ and α is derived from the mean. The interpretation of α and β remains the same.

[4] Note that the SPAR-H method also has an "Insufficient Information" level for each PSF, with a corresponding multiplier of 1. This is not included in Table 1 because Bayesian methods use prior information to enable inference when there is missing information. See [20] for more information.

Table 1: SPAR-H PSFs, levels for each PSF, and multipliers for each level.

PSF	PSF Level	Multiplier
Available Time	Expansive	0.01
	Extra	0.1
	Nominal	1
	Barely adequate	10
	Inadequate	HEP=1.0
Stressors	Nominal	1
	High	2
	Extreme	5
Complexity	Nominal	1
	Moderate	2
	High	5
Experience/Training	High	0.5
	Nominal	1
	Low	3
Procedures	Nominal	1
	Avail., but poor	5
	Incomplete	20
	Not available	50
Ergonomics/HMI	Good	0.5
	Nominal	1
	Poor	10
	Missing/Misleading	50
Fitness for duty	Nominal	1
	Degraded Fitness	5
	Unfit	HEP=1.0
Work Processes	Good	0.5
	Nominal	1
	Poor	5

number of negative PSFs (any PSF where the assigned level has a multiplier greater than 1). Equation 7 is used to calculate the HEP for situations with fewer than three negative PSFs. Equation 8 is used if there are three or more negative PSFs,

$$HEP = NHEP \cdot \prod_{1}^{8} S_i \tag{7}$$

$$HEP = \frac{NHEP \cdot \prod_1^8 S_i}{NHEP \cdot (\prod_1^8 S_i - 1) + 1} \tag{8}$$

where S_i is the multiplier associated with the assigned level of PSF i. For diagnosis tasks $NHEP = 0.01$ and for action tasks $NHEP = 0.001$. The SPAR-H guidance suggests a lower limit of 1×10^5 for the HEP.

5.2. Halden simulator experiments

In 2010, a team of researchers from the Halden Reactor Project conducted a series of experiments at a training simulator of a PWR plant in the United States [36]. The experiments were conducted as part of the NRC's HRA Empirical Study, which was designed to collect simulator data to benchmark and improve existing HRA methods. These experiments used the PWR simulator to challenge the operators under a variety of different contexts.

In these experiments, operating crews each performed four different simulator scenarios: a basic Steam Generator Tube Rupture (SGTR) (scenario 3), a loss of CCW (component cooling water) and RCP (reactor coolant pump) seal water (scenario 2), and a total loss of feedwater (scenario 1A), followed immediately by a complex SGTR (scenario 1C). Each simulator scenario was designed to correspond to an HFE that would be included in a PRA model. For scenario 1A, the criterion for occurrence of an HFE was "failure to establish feed and bleed within 45 minutes of the reactor trip, given that the crews initiate a manual reactor trip before an automatic reactor trip." For scenario 1C, the criterion for occurrence of an HFE was "failure of crew to isolate the ruptured steam generator and control pressure below the SG PORV setpoint to avoid SG PORV opening." In the experiment, the crew was expected to complete this action within 40 minutes of the start of the SGTR. For scenario 2, the criterion for occurrence of an HFE was "failure of the crews to trip the RCPs and start the Positive Displacement Pump (PDP) to prevent RCP seal LOCA." For scenario 3, the criterion for occurrence of an HFE was "failure of crew to isolate the ruptured steam generator and control pressure below the SG PORV setpoint before SG PORV opening."

Four crews participated in the experiments. For three of the crews the composition was one Shift Manager (SM), one Unit Supervisor (US), one Shift Technical Advisor (STA) and two Reactor Operators (RO). In one crew the SM was not present and instead they had three ROs. Three of the crews participated in each of the four scenarios. A fourth crew only participated in three of the scenarios due to simulator failure. In total, the Halden team collected data on fifteen experimental cases.

Multiple types of information were gathered during and after each of the scenarios. The data included second-by-second simulator logs, audio and video recording of the crew, observations from the experimental team, and observations from plant training experts. After the scenarios, the Halden experimental team conducted crew member interviews and distributed questionnaires to all crew members. The crew member questionnaires included an adapted version of the NRC Simulator Crew Evaluation Form, NUREG-SR1020 rev. 9, Form ES-604-2 [41]. Due to the sensitive nature of these data, only the Halden team had access to the raw data. The Halden team processed the data to remove all identifying details about the crews before releasing any information to other research teams.

The Halden team issued report HWR-981, [36], which summarizes the fifteen experiments. The summary includes detailed descriptions of the scenarios, and both the context and the outcome of each experiment. The summary includes ratings for eleven PSFs: stress, adequacy of time, team dynamics, work processes, communication, scenario complexity, indications of conditions, human-machine interface, training and experience, procedural guidance, and execution complexity. The Halden team rated these PSFs on a four-point scale: nominal or positive, not a driver, negative driver, and main negative driver.

5.3. Extracted data

While the Halden team did not directly collect information about the SPAR-H PSFs, the anonymized Halden data contained enough information to be used to extract the SPAR-H PSFs. The Sandia team used the Halden report [36] and the anonymized simulator crew evaluation forms to assign a level for the SPAR-H PSFs for each of the fifteen experiments.

The data extracted from the fifteen simulator runs are presented in Table 2. Each row represents a single crew's performance in one of the four scenarios. The first column of Table 2 contains the scenario number for each data point. The next eight columns of Table 2 are the levels of the SPAR-H PSFs, as rated by the Sandia team. Together, these eight columns form the context of the performance. Each of these contexts is associated with a letter (A-D) for brevity.

At this point, it is important to differentiate between scenarios and contexts. The scenarios denote the configuration of the experiments conducted by Halden. The contexts are the states of the PSFs associated with each experimental run, or each crew-scenario combination. The context for each scenario is not necessarily identical, because several of the PSFs can vary with different crews. The data set in Table 2 includes four contexts. For this set of data, the context is identical across all of crews for each scenario in the experiment. However, this uniformity will not always be seen in simulator data since several of the PSFs can vary with different crews.

The final column of Table 2 documents whether the crew response to the scenario was an error, according to the criteria discussed in the previous section. In this column, "Yes" denotes than an HFE occurred, "No" denotes that there was not an HFE. The error column for Scenario 1C contains two different outcomes for each crew, due to a difference between the experimental failure criterion and the PRA failure criterion for this scenario. For scenario 1C, the PRA failure criterion is "failure to isolate the ruptured SG and control pressure below the SG PORV setpoint to avoid SG PORV opening"; the experimental failure criterion added a 40 minute time cut-off to the PRA criterion. In the error column in Table 2, the first result represents the PRA failure criterion and the result in parenthesis represents the experimental failure criterion.

5.4. Bayesian updating

The Bayesian updating approach described in Section 4 is implemented on each context, or combination of PSFs. For the eight PSFs, in SPAR-H, there are 19440 possible contexts.

The SPAR-H equations, Equations 7 and 8, deterministically assign an HEP for each combination of PSFs, $P(Error|PSF_1 \cap PSF_2 \cap ... \cap PSF_8)$. Relating this back to Section 4.3, SPAR-H assigns a value for $E(p_0)$ for each of 19440 the possible contexts.

To obtain $E(p_0)$ for each context, we developed a Matlab script implementing Equations 7 and 8, the rule for selecting the appropriate equation, and the lower limit on HEP. We used a second Matlab script to assign the parameters of a CNI prior distribution for each p_0. The parameters of the prior distribution for the four contexts represented in the data are shown in Table 3.

To implement the Bayesian approach we need two pieces of data for each context: x, the number of human failures, and n, the number of opportunities for human failure. The data for each context in Table 2 can be aggregated together to provide values for x and n. The aggregated data for the four contexts is presented in Table 4. Two sets of data are presented for context B, to account for the two sets of failure criteria for Halden scenario 1C. The first set of data is for human errors, as defined by the

Scen.	Time	Stressors	Complex.	Exper.	Procedures	HMI	Fitness	WorkProc	Context	Error?
1A	Extra	Nominal	Moderate	Nominal	Avail., but poor	Nominal	Nominal	Nominal	A	No
1A	Extra	Nominal	Moderate	Nominal	Avail., but poor	Nominal	Nominal	Nominal	A	No
1A	Extra	Nominal	Moderate	Nominal	Avail., but poor	Nominal	Nominal	Nominal	A	No
1A	Extra	Nominal	Moderate	Nominal	Avail., but poor	Nominal	Nominal	Nominal	A	No
1C	Barely adeq.	High	Moderate	Nominal	Avail., but poor	Nominal	Nominal	Nominal	B	No(No)
1C	Barely adeq.	High	Moderate	Nominal	Avail., but poor	Nominal	Nominal	Nominal	B	No(Yes)
1C	Barely adeq.	High	Moderate	Nominal	Avail., but poor	Nominal	Nominal	Nominal	B	No(Yes)
1C	Barely adeq.	High	Moderate	Nominal	Avail., but poor	Nominal	Nominal	Nominal	B	Yes(Yes)
2	Inadequate	High	High	Low	Avail., but poor	Nominal	Nominal	Poor	C	Yes
2	Inadequate	High	High	Low	Avail., but poor	Nominal	Nominal	Poor	C	Yes
2	Inadequate	High	High	Low	Avail., but poor	Nominal	Nominal	Poor	C	Yes
2	Inadequate	High	High	Low	Avail., but poor	Nominal	Nominal	Poor	C	Yes
3	Extra	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	D	No
3	Extra	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	D	No
3	Extra	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	D	No

Table 2: SPAR-H PSF assignments and performance outcomes extracted from the fifteen Halden simulator experiments. See Section 5.3 for description of each column.

Table 3: Expected value and prior distribution for p for each of the four contexts represented in the simulator data. The expected value is the HEP calculated from the SPAR-H method. The expected value was used to fit the constrained non-informative prior distribution in the table.

Context	$E(p_0)$	Prior
A	1.00E-3	$p_0 \sim Beta(0.5, 499.5)$
B	1.67E-1	$p_0 \sim Beta(0.5, 2.4975)$
C	1.00	$p_0 \sim Beta(50000, 0.5)$
D	1.00E-4	$p_0 \sim Beta(0.5, 4999.5)$

Table 4: Aggregated data and posterior distribution for p for the four contexts represented in the simulator data. The relevant data are the number of failures, x, and the number of opportunities for failure, n. For context B, we ran the analysis for both interpretations of the data. The expected value of p_1 is calculated from the posterior distribution.

Context	Data (x/n)	Posterior	$E(p_1)$
A	0/4	$p_1 \sim Beta(0.5, 503.5)$	9.92E-4
B_{PRA}	1/4	$p_1 \sim Beta(1.5, 5.4975)$	2.14E-1
B_{exp}	3/4	$p_1 \sim Beta(3.5, 3.4975)$	5.00E-1
C	4/4	$p_1 \sim Beta(50004, 0.5)$	≈ 1.00
D	0/3	$p_1 \sim Beta(0.5, 5002.5)$	1.00E-4

PRA criteria. The second set of data is for human errors, as defined by the experimental criteria.

To obtain the posterior distribution, we apply Equation 5 to combine the prior and the data. The posterior distribution for p_1 for the four updated contexts is presented in Table 4.

5.5. Results

Table 4 demonstrates that it is possible to use sparse data to refine HEPs for existing HRA methods. Furthermore, comparing the posteriors (Table 4) with the priors (Table 3) shows the importance of incorporating data into existing HRA methods.

For contexts A, C, and D, the posterior mean is very close to the prior mean. This indicates that the data from the experiments was consistent with the predictions of the SPAR-H model. For contexts A, C, and D, combining data with the SPAR-H method doesn't change the HEPs assigned. However, adding the data enhances the technical basis of the SPAR-H HEPs, and this enhanced technical basis increases the credibility of the SPAR-H method.

For context B, the posterior mean is measurably different than the prior mean for both interpretations of the data. In both cases, the HEPs increase. For case B_{PRA}, the posterior mean is a slight increase over the prior mean. For case B_{exp}, the posterior mean increases more substantially. This increase shows that the HEPs predicted by SPAR-H underestimate the probability of human errors, which leads to non-conservative PRA results.

The results for context B show that simulator data can be used to refine the probabilities from existing HRA methods. In this case, adding data to the SPAR-H method has both demonstrated and improved a weakness in the original method. Adding data brings the SPAR-H HEPs more in line with the data that is observed. In all cases, adding data increases the credibility of the SPAR-H model.

6. Discussion

The results of the case study demonstrate that Bayesian updating can be used to refine the probability assignments from existing HRA methods. The results also show that the SPAR-H method underestimates the human contribution to risk for at least one context, which leads to non-conservative PRA results. This underestimation provides a powerful argument for the need to use data to refine existing HRA methods.

6.1. Applicability of simulator data

Simulator data provides powerful insight into $P(Error|PSFs)$. Leveraging simulators provides a unique opportunity to gather, via controlled experiments, data for events that would not normally be seen in the historical record for NPP operation (e.g., severe accidents, or unlikely PSF states such as unavailable procedures). These controlled experiments can be designed to focus on specific causal factors in order to make inference on performance under known conditions.

It can be argued that simulator data does not always represent the expected conditions in a nuclear power plant (e.g., the simulator events are designed to be more difficult than "real" accidents). However, the approach presented in this paper bypasses this difference, by reasoning about $P(Error|PSFs)$ rather than $P(Error)$. While the simulator data is not necessarily perfectly representative of the context of real operational events, the simulator data is fully representative of the occurrence of human failure events, given the PSFs. That is, we believe that operators responding to a given context will exhibit the same response in both real events and in simulator performance.

6.2. Challenges: data access and organization

The greatest challenge experienced was gaining access to the data. The most relevant information comes from a wide range of simulator data (several sources of such data are described in Section 4.2) and operational experience (including HERA [42], OPERA [38], and CORE-DATA [43]). These databases provide a powerful opportunity to update existing HRA methods. However, access to the data is severely limited due to the sensitivity of the raw data.

The effort taken by the Halden reactor project to anonymize the raw data resulted in an extremely rich set of information. To truly enable use of data for HRA, future data collection projects should follow Halden's lead by anonymizing the data and making it available to research organizations.

The simulator experiments provide a good starting point for data collection. It is imperative to continue obtaining simulator data. However, the scope of HRA data must be expanded to include specific human-performance elements (PSFs) in the operational data. Since HRA models have a variety of PSFs, the data required to update existing HRA methods must span the spectrum of PSFs used in current HRA methods. The PSF taxonomy proposed by [44] provides a good starting point for expanding operational data collection frameworks to include a comprehensive set of PSFs.

On a related note, there is a broad range of information, beyond simulator and operational data, that is relevant to human performance in nuclear power plants. Possible sources of information that may be relevant to HRA hypotheses include: existing HRA methods, expert judgment, and cognitive literature. These information sources provide a variety of both qualitative and quantitative information, including point and interval estimates of HEPs, linear models, correlations between various PIFs, the magnitude of PSF influences, and relationships between variables. The data from these sources may be less sensitive than the simulator and operational data. However, no one has made a concerted effort to develop a database containing this information. Effort should be undertaken to assemble this information in a database accessible to the HRA research community.

6.3. Next steps

The case study in this paper only included data about four of the contexts in SPAR-H. As a next step, it is desirable to continue to gather simulator data, ideally from multiple simulators, to capture the other contexts in the SPAR-H method.

In Section 5.4, we presented results for two different interpretations of the Halden data (contexts B_{PRA} and B_{exp}). In a more complicated Bayesian analysis, we could directly address the uncertainty in the data by assigning prior distributions to the number of failures in the data.

For this case study, we used a CNI prior distribution. The CNI distribution is known to have relatively light tails, which makes it difficult to update with sparse data for low probability events. In future work, other diffuse priors (including mixture priors) should be compared to the CNI prior.

7. Conclusions

This paper presented a Bayesian methodology to update HEPs from existing HRA methods with simulator data. We provided a case study, wherein we updated several values from SPAR-H with simulator data from Halden experiments. This methodology and case study demonstrate the value of Bayesian methods for HRA. Furthermore, we demonstrated that we do not require an inordinate amount of data, nor do we require perfect data to improve existing HRA approaches. The same Bayesian approach can be used to add operational data to HRA methods, if the operational data includes descriptions of the PSFs that are relevant to the operational performance.

Furthermore, by using the Bayesian Network framework, it is possible to combine data about $P(Error|PSFs)$ with data about $P(PSFs)$ to provide an expanded technical basis for HRA, with the added benefit of an expanded scope of reasoning. The BN version of the SPAR-H model, documented in [20] can be updated using the same methodology documented in this paper. Using the Bayesian updating on the probabilities produced by BN models provides the ability to incorporate a wide variety of data sources to enhance the credibility of the model.

The Bayesian inference process can be applied to a wide range of HRA problems. In this paper, we demonstrated how to conduct inference for HRA problems at a high level of abstraction (failure probabilities for general human tasks). Future works should be developed to demonstrate how the Bayesian inference process can be used to conduct inference on HRA problems at more detailed, causal levels (e.g., what is the effect of a given PSF on performance).

Implementing the Bayesian methodology brings HRA in line with the Bayesian processes used in other aspects of PRA. By expending effort to bridge the gap between hardware and human failure modeling, HRA quality (and thereby PRA quality) is enhanced. Inherent deficiencies in any part of PRA degrade the applicability of the PRA results. The Bayesian approach provides a way to incorporate data into the HRA process, which will go a long way to resolve the model credibility issues that plague HRA.

8. Acknowledgments

The authors thank the Halden Reactor Project for providing the anonymized experimental data and for reviewing this manuscript. This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

9. References

[1] Committee on the Safety of Nuclear Installations (CSNI), . Simulator studies for HRA purposes. NEA/CSNI/R(2012)1; OECD Nuclear Energy Agency; 2012.

[2] Aven, T.. Foundations of risk analysis. John Wiley & Sons; 2003.

[3] Siu, N.O., Kelly, D.L.. Bayesian parameter estimation in probabilistic risk assessment. Reliability Engineering & System Safety 1998;62(1):89–116.

[4] Atwood, C.L., LaChance, J.L., Martz, H.F., Anderson, D.J., Englehardt, M., Whitehead, D., et al. Handbook of parameter estimation for probabilistic risk assessment. NUREG/CR-6823; US Nuclear Regulatory Commission; Washington DC; 2003.

[5] Kelly, D.L., Smith, C.L.. Bayesian inference in probabilistic risk assessment–the current state of the art. Reliability Engineering & System Safety 2009;94(2):628–643.

[6] Kelly, D., Smith, C.. Bayesian Inference for Probabilistic Risk Assessment: A Practitioner's Guidebook. Springer; 2011.

[7] French, S., Bedford, T., Pollard, S.J.T., Soane, E.. Human reliability analysis: A critique and review for managers. Safety Science 2011;49(6):753–763.

[8] Hildebrandt, M., Oxstrand, J., Boring, R.L.. Living HRA: Building new communities of practice for proactive safety management. In: Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting; vol. 53. Human Factors and Ergonomics Society; 2009, p. 1579–1583.

[9] Adhikari, S., Bayley, C., Bedford, T., Busby, J.S., Cliffe, A., Devgun, G., et al. Human reliability analysis: A review and critique. Manchester Business School Working Paper No. 589; Manchester Business School, The University of Manchester; Manchester, UK; 2009. URL `http://www.mbs.ac.uk/research/workingpapers/`.

[10] Committee on the Safety of Nuclear Installations (CSNI) Working Group on Risk (WGRisk), . HRA data and recommended actions to support the collection and exchange of HRA data. NEA/CSNI/R(2008)9; OECD Nuclear Energy Agency; 2008.

[11] Mosleh, A., Chang, Y.H.. Model-based human reliability analysis: prospects and requirements. Reliability Engineering and System Safety 2004;83(2):241–253.

[12] Dougherty, E.M.. Human reliability analysis–where shouldst thou turn? Reliability Engineering & System Safety 1990;29(3):283–299.

[13] Kim, M.C., Seong, P.H., Hollnagel, E.. A probabilistic approach for determining the control mode in CREAM. Reliability Engineering & System Safety 2006;91(2):191–199.

[14] Trucco, P., Cagno, E., Ruggeri, F., Grande, O.. A Bayesian belief network modelling of organisational factors in risk analysis: A case study in maritime transportation. Reliability Engineering & System Safety 2008;93(6):845–856.

[15] Groth, K.M.. A data-informed model of performance shaping factors for use in human reliability analysis. Ph.D. thesis; University of Maryland; College Park, MD; 2009. URL `http://gradworks.umi.com/33/91/3391233.html`.

[16] Baraldi, P., Conti, M., Librizzi, M., Zio, E., Podofillini, L., Dang, V.N.. A Bayesian network model for dependence assessment in Human Reliability Analysis. In: Proceedings of the European Society for Reliability Annual Meeting (ESREL 2009). Prague; 2009,.

[17] Wang, Y.F., Xie, M., Ng, K.M., Habibullah, M.S.. Probability analysis of offshore fire by incorporating human and organizational factor. Ocean Engineering 2011;38(1718):2042–2055.

[18] Groth, K.M., Mosleh, A.. Deriving causal Bayesian networks from human reliability analysis data: A methodology and example model. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2012;226(4):361–379.

[19] Mindock, J.. Development and application of spaceflight performance sharping factors for human reliability analysis. Ph.D. thesis; University of Colorado; Boulder CO; 2012.

[20] Groth, K.M., Swiler, L.P.. Bridging the gap between HRA research and HRA practice: A Bayesian network version of SPAR-H. Reliability Engineering and System Safety 2013 (In Press, Accepted Manuscript);doi: 10.1016/j.ress.2013.02.015.

[21] Sundaramurthi, R., Smidts, C.. Human reliability modeling for the Next Generation System Code. Annals of Nuclear Energy 2013;52:137–156.

[22] Martins, M.R., Maturana, M.C.. Application of bayesian belief networks to the human reliability analysis of an oil tanker operation focusing on collision accidents. Reliability Engineering and System Safety 2013;110:89–109.

[23] Hallbert, B., Kolaczkowski, A.. The employment of empirical data and Bayesian methods in human reliability analysis: A feasibility study. NUREG/CR-6949; US Nuclear Regulatory Commission; Washington DC; 2007.

[24] Chang, Y.H.J., Mosleh, A., Roth, E.M., Richards, R., Shen, S.H., Bley, D., et al. Methodology for collection and analysis of simulator data for HRA applications. In: Proceedings of the International Conference on Probabilistic Safety Assessment and Management (PSAM 11). Helsinki, Finland; 2012,.

[25] Swain, A.D.. Accident sequence evaluation program human reliability analysis procedure. NUREG/CR-4772; US Nuclear Regulatory Commission; 1987.

[26] Parry, G.W., N., B.A., J., S.A., P., M.. An approach to the analysis of operator actions in probabilistic risk assessment. EPRI TR-100259; The Electric Power Research Institute (EPRI); 1992.

[27] Reer, B., Dang, V.N.. The commission errors search and assessment (CESA) method. PSI Report Nr. 07-03; Paul Scherrer Institut; Switzerland; 2007.

[28] Hollnagel, E.. Cognitive Reliability and Error Analysis Method (CREAM). New York: Elsevier; 1998.

[29] Williams, J.C.. HEART–a proposed method for assessing and reducing human error. In: 9th Advances in Reliability Technology Symposium, University of Bradford. 1986,.

[30] Jung, W., et al. A standard HRA method for PSA in nuclear power plant; K-HRA method,. KAERI/TR-2961/2005; Korean Atomic Energy Research Institute (KAERI); 2005.

[31] Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B., Rea, K.. SLIM-MAUD: an approach to assessing human error probabilities using structured expert judgement. NUREG/CR-3518; US Nuclear Regulatory Commission; 1984.

[32] Gertman, D., Blackman, H., Marble, J., Byers, J., Smith, C.. The SPAR-H Human Reliability Analysis method. NUREG/CR-6883; US Nuclear Regulatory Commission; Washington DC; 2005.

[33] Swain, A.D., Guttmann, H.E.. Handbook of Human Reliability Analysis with emphasis on nuclear power plant applications. NUREG/CR-1278; US Nuclear Regulatory Commission; Washington DC; 1983.

[34] Cooper, S.E., Ramey-Smith, A.M., Wreathall, J., Parry, G.W., Bley, D.C., Luckas, W.J., et al. A technique for human error analysis ATHEANA. NUREG/CR-6350; US Nuclear Regulatory Commission; Washington DC; 1996.

[35] Simulator-based Human Factors Studies Across 25 Years: The History of the Halden Man-Machine Laboratory. Springer; 2010.

[36] Broberg, H., Hildebrandt, M., Nowell, R.. Results from the 2010 HRA data collection at a US PWR training simulator. HWR-981; OECD Halden Reactor Project; Halden, Norway; 2011.

[37] Park, J., Jung, W.. A database for human performance under simulated emergencies of nuclear power plants. Nuclear Engineering and Technology 2005;37(5):491.

[38] Park, J., Jung, W.. OPERA–a human performance database under simulated emergencies of nuclear power plants. Reliability Engineering & System Safety 2007;92(4):503–519.

[39] Chang, Y.H.J., Lois, E.. Overview of the NRC's HRA data program and recent activities. In: Proceedings of the International Conference on Probabilistic Safety Assessment and Management (PSAM 11). Helsinki, Finland; 2012,.

[40] Atwood, C.L.. Constrained noninformative priors in risk assessment. Reliability Engineering and System Safety 1996;53(1):37–46.

[41] Form ES-604-2: Simulator crew evaluation form. NUREG-1021 rev. 9; US Nuclear Regulatory Commission; 2007.

[42] Hallbert, B., Boring, R., Gertman, D., Dudenhoeffer, D., Whaley, A., Marble, J., et al. Human events repository analysis (HERA) system overview. NUREG/CR-6903, Vol. 1; US Nuclear Regulatory Commission; Washington DC; 2006.

[43] Kirwan, B., Basra, G., Taylor-Adams, S.. CORE-DATA: a computerised human error database for human reliability support. In: Proceedings of the 1997

IEEE Sixth Conference on Human Factors and Power Plants. Orlando, FL; 1997, p. 9.7–9.12. Global Perspectives of Human Factors in Power Generation.
[44] Groth, K.M., Mosleh, A.. A data-informed PIF hierarchy for model-based human reliability analysis. Reliability Engineering & System Safety 2012;108:154–174.